5) >> clear all

>> close all

>> rng(0,'twister');

>> a = -2;

>> b = 2;

>> r = (b-a).\*rand(20,1) + a;

>> r\_range = [min(r) max(r)]

r\_range =

-1.6098 1.8824

>> p = [0.3 -0.6 0.05 -3];

>> v = polyval( p, r );

>> R = normrnd(v,0.25)

R =

-3.1215

-3.5186

-5.2275

-2.7940

-2.9750

-5.6284

-3.5422

-3.0856

-3.0058

-3.2495

-4.7424

-3.3177

-3.3477

-3.2074

-4.0218

-4.8250

-3.0023

-3.3852

-2.9396

-3.5002

>> plot( r, R, 'o' )

>> [data3, gof3] = fit( r, R, 'poly3' );

>> [data5, gof5] = fit( r, R, 'poly5' );

>> hold on

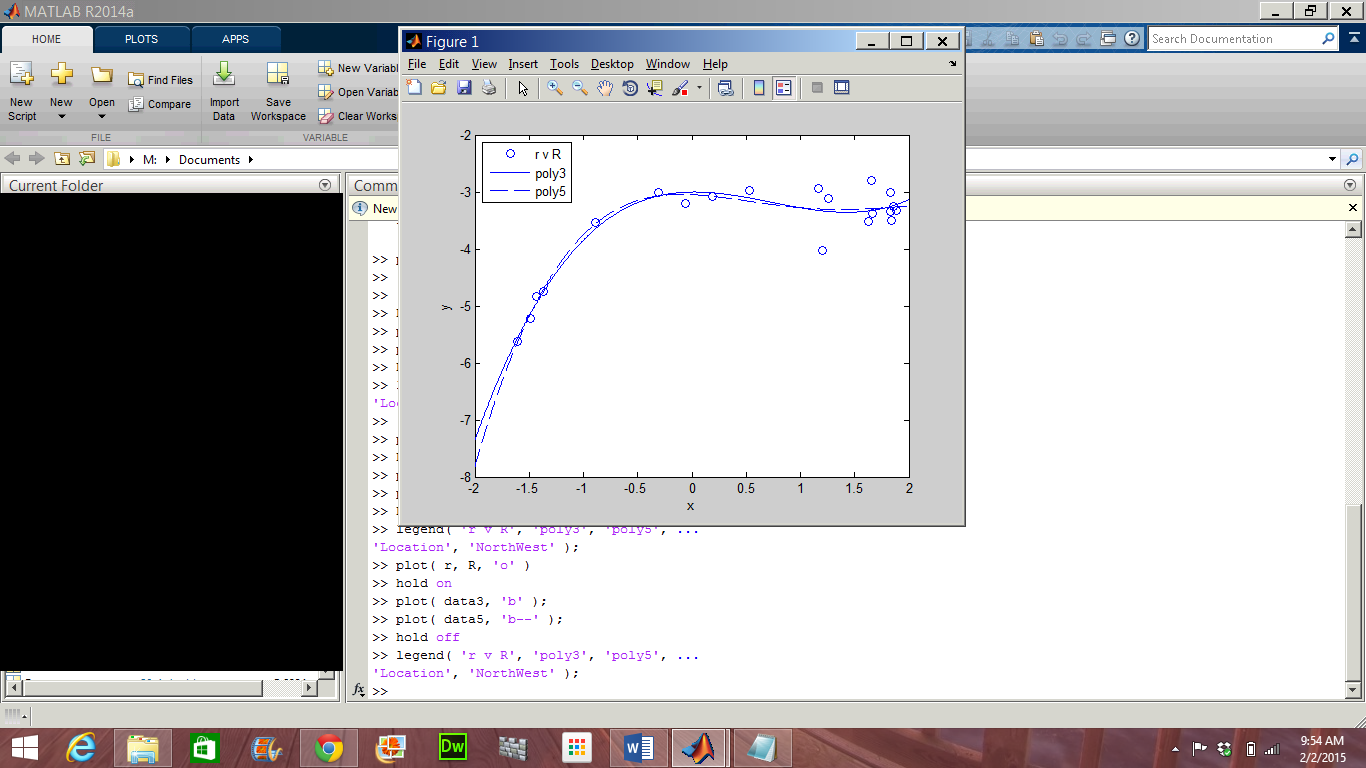
>> plot( data3, 'b' );

>> plot( data5, 'b--' );

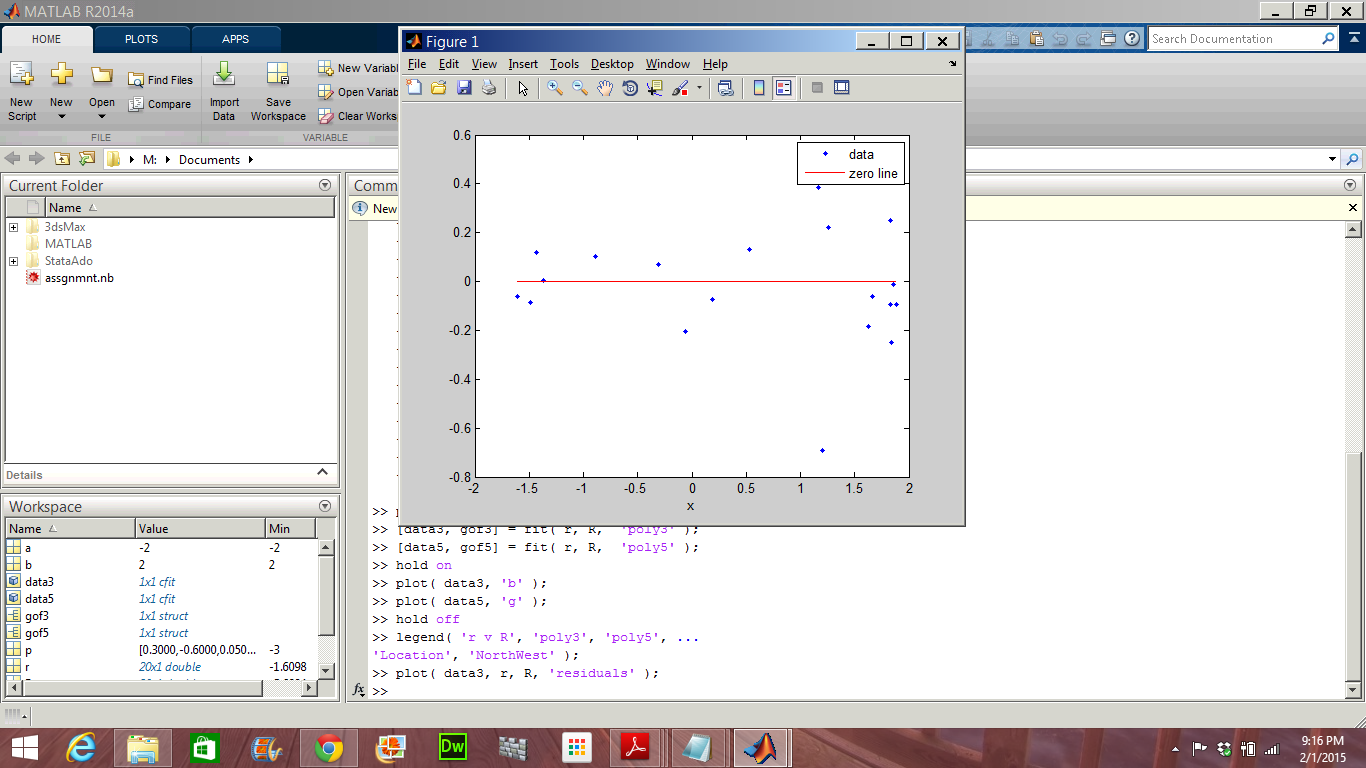
>> hold off

>> legend( 'r v R', 'poly3', 'poly5', ...

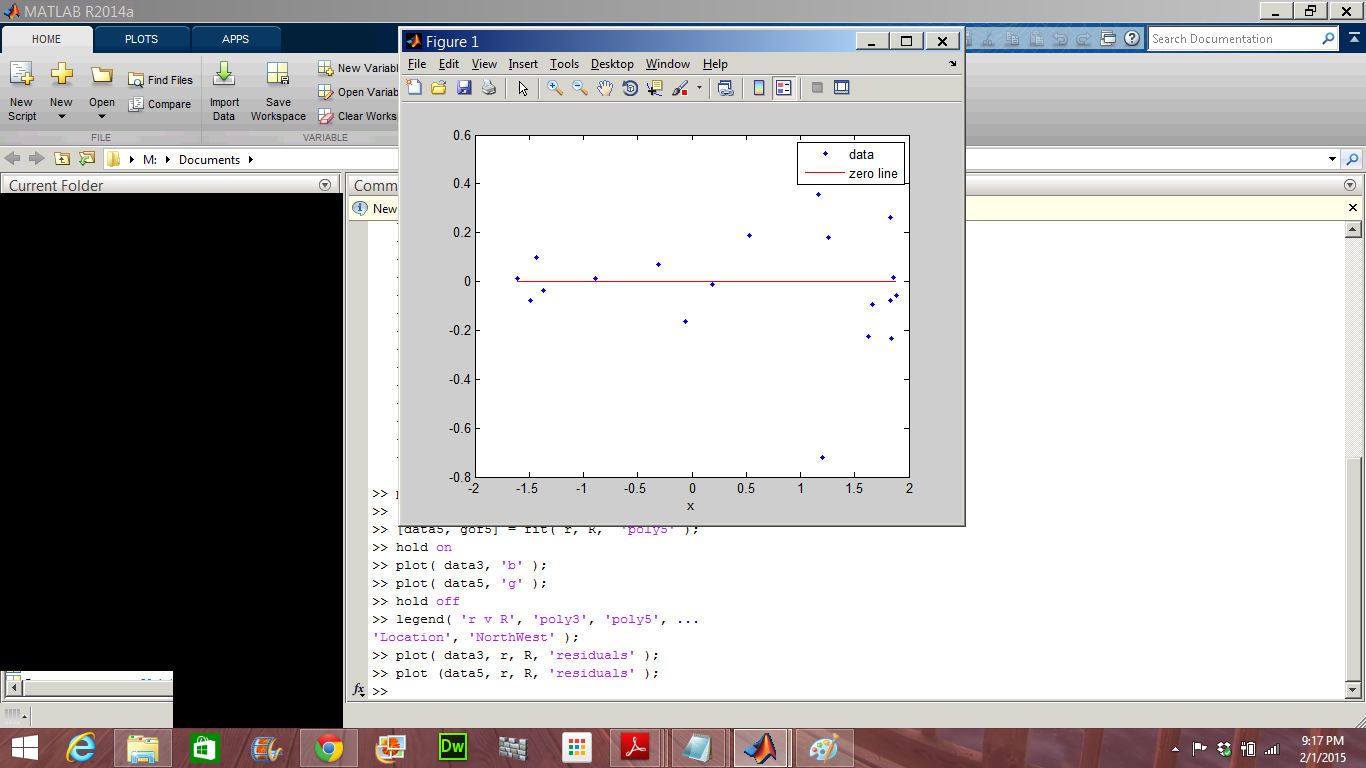
'Location', 'NorthWest' );



>> plot( data3, r, R, ‘residuals’);



>> plot( data5, r, R, ‘residuals’);



The residuals and fits for both the polynomials are similar. Now, we will examine the behavior of the fits for a different range of x.

>> plot( r, R, 'o' )

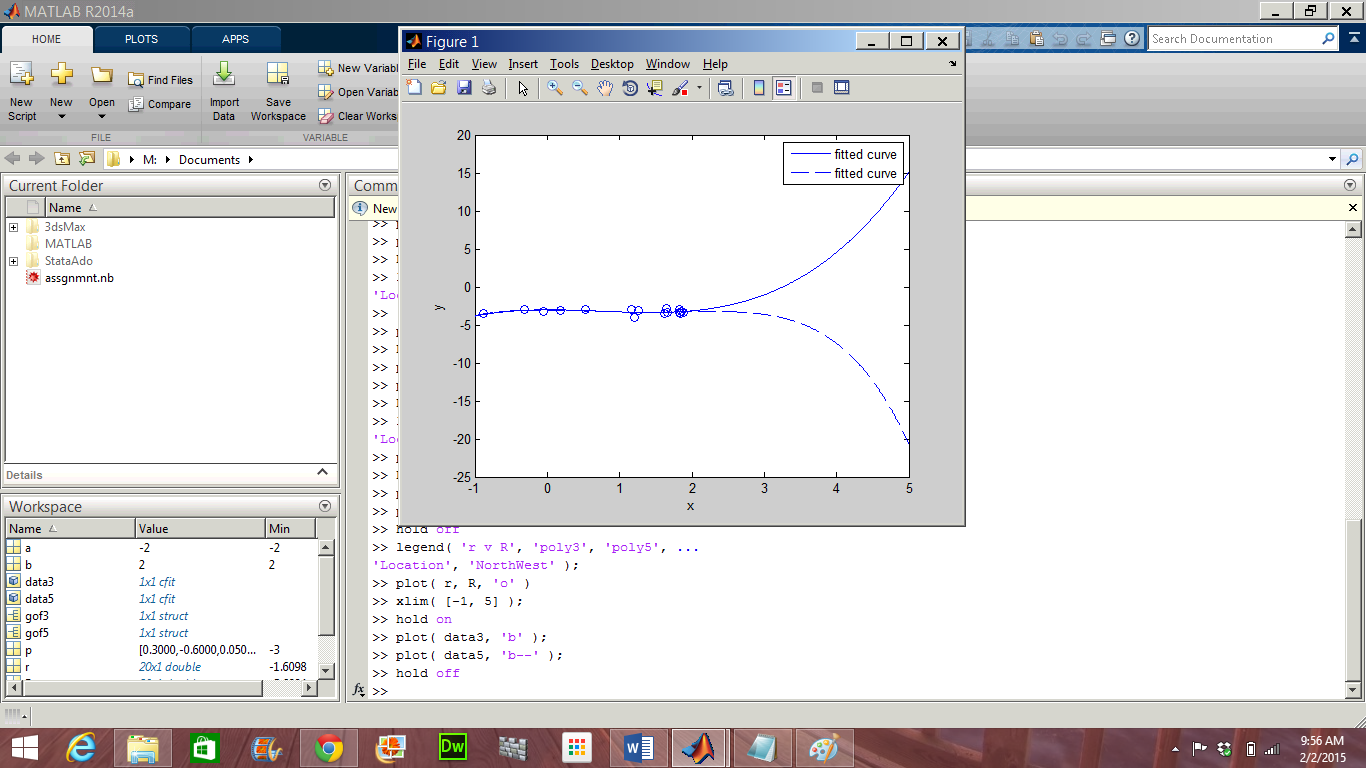
>> xlim( [-1, 5] );

>> hold on

>> plot( data3, 'b' );

>> plot( data5, 'b--' );

>> hold off



gof provides us with goodness-of-fit statistics.

>> gof3

gof3 =

sse: 1.2409

rsquare: 0.9038

dfe: 16

adjrsquare: 0.8858

rmse: 0.2785

>> gof5

gof5 =

sse: 1.2063

rsquare: 0.9065

dfe: 14

adjrsquare: 0.8731

rmse: 0.2935

The sse statistic is the least square error of the fit, with a value closer to 0 indicating a better fit. The lower sse value is associated with the 5th degree polynomial. However, the behavior of this fit beyond the data range makes it a poor choice for extrapolation.

We can get the confidence intervals by:

>> ci = confint( data3 )

ci =

0.1000 -0.7128 -0.3576 -3.2772

0.4125 -0.4078 0.4147 -2.7234

>> ci = confint ( data5 )

ci =

-0.2291 -0.3113 -0.5352 -1.1006 -0.9285 -3.3941

0.2191 0.2084 1.1875 0.2701 0.7619 -2.6987

The confidence bounds on the coefficients determine their accuracy. The bounds cross zero on several coefficients for the fifth-degree polynomial. This means we cannot be sure that these coefficients differ from zero. If the higher order model terms may have coefficients of zero, they are not helping with the fit, which suggests that this model over fits the data. Therefore, after examining both the graphical and numerical fit results, we should select the 3rd degree polynomial as the best fit to extrapolate the data.